

# Rozvoj otevřených archivů elektronických tisků a alternativních bibliografických služeb v oboru informační věda

Eva Bratková

*Ústav informačních studií a knihovnictví FF UK v Praze*

Eva.Bratkova@ff.cuni.cz

## 1 Tradiční a novodobé informační zabezpečení oboru informační věda

Obor **informační věda** je v současné době reprezentován již relativně velkým množstvím literatury z hlediska časového i prostorového. Informační potřeby odborníků jsou standardně satureovány, tak jako v jiných oborech, informačními zdroji a službami různého charakteru. Bibliografické potřeby jsou zajišťovány prestižními bibliografickými databázemi oboru informační věda. Jde především o trojici známých mezinárodních bází: ISTA (Information Science and Technology Abstracts) produkovanou pro ASIST (The American Society for Information Science and Technology) společností Information Today, LISA (Library and Information Science Abstracts) produkovanou v současnosti společností Cambridge Scientific Abstracts a LLIS (Library Literature and Information Science) produkovanou společností H.W. Wilson. Jde o zdroje americké provenience, které určitě důkladně mapují početnou literaturu Severní Ameriky, pokrytí literatury evropské provenience, zejména pak východoevropské, má ale bílá místa. K těmto zdrojům je potřeba zařadit i části významných mezinárodních databází polytematického charakteru, jakými jsou báze PASCAL s třídou 001A01 (vědy o informacích a dokumentace) francouzského institutu INIST nebo báze VINITI se sekci 59 (informatika) stejnojmenného institutu. Důležitou roli hrají rovněž oborové báze národního charakteru, pokud existují. Návazný systém služeb tradičního nebo elektronického charakteru zabezpečuje odborníkům úplné texty požadovaných dokumentů. Převážná část těchto produktů a služeb je ovšem komerčního charakteru. Pokud na ně chybějí prostředky, nastává pro odborníky či badatele jistý problém.

Obor informační věda je však dnes prezentován také stále narůstajícím množstvím volně dostupné literatury a dalších informačních zdrojů v elektronické formě uložených na serverech tzv. „povrchového“ webu nebo v databázích či repozitářích dostupných v rámci tzv. webu „hlubokého“ či „neviditelného“. Celosvětově přibývá volně dostupných elektronických časopisů i z našeho oboru, existuje již nepřehledné množství volně dostupných konferenčních sítel s kvalitními sborníky úplných textů příspěvků, sídla výzkumů s obsažnými výzkumnými a technickými zprávami, volně k dispozici jsou na řadě serverů různých organizací oboru další materiály legislativní, výukové, instruktážní a jiné povahy. A právě tento volně dostupný materiál je předmětem zájmu různých nově vznikajících systémů **digitálních knihoven**, informačních bran, portálů atd., které operují zpravidla na mezinárodní nebo národních úrovních. Specifické místo mezi nimi zaujímají tzv. **archivy elektronických tisků**. Dá se říci, že ve hře je možná i něco více než jen volně dostupný dokument. Ve sféře vědy, realizované zejména v oblasti akademicko-univerzitní začal, v jistém slova smyslu, „boj“ o autory mezi komplexem tradiční komerční vydavatelské a distributorské sféry a alternativními novodobými systémy otevřených archivů elektronických tisků a otevřených digitálních knihoven a jejich návazných služeb. Jinak řečeno, některé novodobé systémy, těžící z moderních technologií, zaujaly doposud téměř prázdný prostor mezi sférou autorů a sférou nakladatelů. To vše v zájmu uživatelů informací také z oboru informační věda, kteří stále, každým dnem, narážejí na nejrůznější bariéry při získávání potřebných a požadovaných informací k dalšímu rozvíjení tohoto oboru a taktéž návazné praktické činnosti.

I v České republice již existuje velké množství literatury a dalších volně dostupných zdrojů z oboru informační věda v elektronické formě, jsou však rozptýleny na serverech různých informačních, knihovnických, vzdělávacích a jiných organizací. Jejich komplexní registrace na oborové úrovni neexistuje. K dispozici u nás není prozatím ani žádný elektronický archiv, který by tyto dokumenty na oborové úrovni archivoval a poskytoval tak adekvátní služby jejich uživatelům. Dá se hypoteticky předpokládat, že ani obsah nevelké množiny elektronických volně dostupných časopisů české a popřípadě i sousední slovenské provenience není v optimálním množství sledován a využíván ze strany odborníků (nemluvě o časopisech zahraničních). Je možná na čase uvažovat o možnosti vybudování takového systému. Měl by zároveň poskytovat záruky na mezinárodní propojení s jinými již existujícími systémy a jejich službami. Bylo by dobré více zviditelnit kvalitní české dokumenty vůči světu zapojením se do sítě již provozovaných archivů elektronických tisků a digitálních knihoven v oboru informační věda. V principu jde buď o ukládání významných dokumentů do již existujících zahraničních archivů elektronických tisků (minimálně na evropské úrovni), nebo alespoň o nabídnutí kvalitních metadat o našich dokumentech z oboru informační věda směrem do systémů digitálních knihoven. Je potřebné počítat ale s tím, že kvalitní metadata znamenají minimálně abstrakt a popřípadě i název dokumentu v anglickém jazyce, pokud bude dokument vytvořen v jazyce národním. Některé kvalitní časopisy tyto údaje mají (například Národní knihovna : knihovnická revue), a proto je velká škoda, že jich není využito v zahraničních službách (bohužel, citovaný časopis například není vůbec excerpován ani jednou z výše zmiňovaných komerčních světových bibliografických služeb). Nabízejí se však novodobé bibliografické služby jako alternativa služeb komerčních. I o nich je následující text článku.

## 2 Otevřené archivy elektronických tisků ve vědě a jejich programové zabezpečení

Komunikace informací ve vědě doznala díky rozvoji informačních technologií v průběhu posledních 13 let značných proměn. Jedním z výrazných trendů je další rozvoj a rozšiřování „přímé vědecké komunikace“ (fr. communication scientifique directe) pomocí **otevřených archivů elektronických tisků** (Open Eprint Archives). Jsou jimi míněny, jednoduše řečeno, systémy zahrnující fondy archivovaných digitálních dokumentů [Eprints.org, 2002, What is an Eprint Archive]. U současných archivů elektronických tisků se zdůrazňuje jejich otevřenost v intencích principů a protokolů Iniciativy otevřených archivů OAI (Open Archives Initiative, <http://www.openarchives.org>). Otevřeností je samozřejmě míněn volný přístup k fondům digitálních informací ze strany koncových uživatelů, především je jí však míněn volný přístup ze strany libovolných informačních systémů (včetně internetových vyhledávačů), které pomocí specifických programů (harvesterů) mohou volně „**sklízet metadata**“ (metadata harvesting), aby na základě nich mohly poskytovat nadstavbové komplexní služby. Doporučovaným standardem pro volnou komunikaci metadat od poskytovatelů metadat k poskytovatelům služeb je v současné chvíli protokol pro sklizeň metadat **OAI-PMH** (Protocol for Metadata Harvesting) ve své stabilní verzi 2.0 [LAGOZE, 2003; CARPENTER, 2001], k jehož aplikaci přistoupilo již velké množství systémů, další se připravují. Standard je rozvíjen iniciativou OAI. Archivy, které již aplikovaly protokol OAI-PMH, jsou označovány za OAI vyhovující archivy [Eprints.org, 2002, What is OAI-compliance]. Příklad reálného metadatového záznamu z archivu elektronických tisků, získaného na základě protokolu OAI-PMH je vidět v další třetí části na obr. č. 4.

Specifikem archivů elektronických tisků je typ archivovaného materiálu. **Elektronickým tiskem** je míněn digitální text recenzovaného vědeckého článku, a to v době před nebo po jeho recenzování a publikování (zpravidla v časopise nebo konferenčním sborníku). Může

tedy jít o tzv. „**preprint**“ (první návrh textu článku) nebo tzv. „**postprint**“ (recenzovaný a publikovaný poslední návrh textu článku). Elektronickým tiskem je rovněž míněna jakákoliv další verze textu článku realizovaná v libovolné fázi mezi dvěma uvedenými krajními polohami, jakož i jakákoliv další aktualizace již publikovaného článku [Eprints.org, 2002, What is an Eprint]. Nové verze (všechny se postupně archivují) mohou ve značné míře vznikat na základě odezvy a připomínkování uživatelů a lze tak hovořit o novodobé formě recenzní činnosti. V rámci archivů se rozvíjí i novodobé citování na základě identifikátorů elektronických tisků. Rozvíjejí se také automatické systémy tvorby citačních indexů nad archivy úplných dokumentů. Archiv elektronických tisků budovaný několik let podává unikátní svědectví o rozvoji dané disciplíny v čase i prostoru. Pokud se do archivů ukládají i jiné typy informačních materiálů než jenom elektronické tisky, může se pro ně v praxi uplatňovat již širší pojmenování jako „**elektronický archiv**“ nebo „**digitální knihovna**“ aj.

Specifickou otázkou archivů elektronických tisků je **autorskoprávní problematika**. Již výše zmíněný prostor mezi autorem a vydavatelem lze označit téměř za „neuralgický“, v rámci kterého se odehrává nejedna diskuse a nejedno jednání ohledně autorských práv. Pokud jde o preprinty, je v rámci iniciativy volně dostupných archivů rozšiřován názor, že právo copyrightu drží autor, který vlastně nemusí nikoho žádat o povolení uložit text do volně dostupného archivu. Problémy mohou nastat v momentu, kdy autor chce dokument předložit komerčnímu vydavateli. U těchto dokumentů, stejně jako těch, které již postoupily do recenzního řízení nebo které byly již publikovány, musí nastat jednání s vydavatelem o podmínkách možného uložení dokumentu do volně dostupného archivu.

Dalším výrazným specifikem archivů elektronických tisků je tzv. „**auto-archivace**“ digitálních textů, tedy jejich samostatné uložení do repozitáře archivu ze strany autora (angl. self-archiving), a také jejich bibliografické zpracování pomocí interaktivního formuláře pro tvorbu **metadat**. Jejich kontrola je zpravidla náplní práce editorů archivů, zejména pokud jde o údaje věcného charakteru.

Archivy mohou být rozdělovány z hlediska místa vzniku dokumentu na archivy lokální, národní či mezinárodní a z hlediska obsahu dokumentu na archivy univerzální, víceoborové nebo oborové či tematické povahy. Archivy mohou být budovány jako centralizované (s jediným repozitářem) nebo distribuované či „virtuální“ (s více repozitáři a centrální databází metadat včetně jednotného rozhraní pro vyhledávání). Distribuované systémy jsou pak často označovány termínem „digitální knihovna“. K nejvýznamnějším mezinárodním archivům elektronických tisků s centralizovaným repozitářem patří například archiv pro fyziku a jiné exaktní vědy **arXiv.org** (<http://arXiv.org>) a archiv pro kognitivní vědy **CogPrints** (<http://cogprints.soton.ac.uk>). K zástupcům digitálních knihoven mezinárodního rozsahu s distribuovanými repozitáři patří zejména digitální knihovna pro počítačovou vědu **NCSTRL** (<http://www.ncstrl.org>) a digitální knihovna disertací a diplomových prací **NDLTD** (<http://www.ndltd.org>). Uvedené systémy vznikaly již od počátku 90. let minulého století. Bylo jim věnováno mnoho odborných publikací v zahraničním i domácím tisku [viz např. BRATKOVÁ, 2000, 2002].

Významným rysem archivů elektronických tisků je lepší **viditelnost**, **dostupnost**, **využívání** a **vliv** uložených dokumentů na další a rychlý rozvoj disciplín. Po roce 2000 dochází celosvětově k pozvolně se zvyšujícímu nárůstu archivů elektronických tisků a stejně tak počtu dokumentů v nich uložených. Trendem je rozvoj archivů budovaných buď na oborové bázi, nebo na bázi víceoborové. Archivy mohou být budovány i například v rámci nějaké univerzity zahrnující více oborů. Jedním z těchto oborů je i informační věda (včetně knihovnictví, komunikace, informačních technologií, archivnictví, vydavatelství, muzejnictví aj.). Text tohoto článku je věnován právě těmto archivům a bude sledovat i jejich různorodé profily, pokud jde o průniky k dalším disciplínám. Dalším směrem, který bude sledován, je integrace

různorodých zdrojů z oboru informační věda do jednotné informační služby s výraznou novodobou bibliografickou funkcí (viz systém DoIS ve 4. části).

Archivy z oblasti informační věda využívají, tak jako archivy z jiných oblastí, již zavedené aplikační programy. Patří k nim například program „**DSpace**“ (<http://www.dspace.org/>), vyvinutý na Technologickém institutu v Massachusetts (MIT, USA), zejména však v Evropě dobře známý program „**Eprints.org**“ (<http://software.eprints.org/>). Na konci roku 2003 byl instalován v 113 archivech (95 ve verzi 2 a 18 ve verzi 1). Program Eprints.org je aplikován u všech dále v textu analyzovaných archivů z oblasti informační vědy, proto je vhodné uvést o něm několik základních informací. Je zdarma dostupný (free software) pro nekomerční využití (GNU GPL, General Public Licence) a patří k tzv. otevřeným zdrojům (Open Sources). Program je tedy možné upravovat a rozšiřovat. Vznikl na fakultě počítačové vědy na Univerzitě v Southamptonu (Velká Británie) a jeho autorem je Robert Tansley (verze 1). Je novodobou verzí softwaru, který vznikl již v roce 1991 a je doposud využíván výše zmíněným archivem arXiv.org a který později adaptoval také archiv CogPrints. Rozvoj programu je finančně podporován britskou Společnou komisí pro informační systémy JISC (The Joint Information Systems Committee, <http://www.jisc.ac.uk/>) a americkou Národní vědeckou nadací NSF (The National Science Foundation, National, <http://www.nsf.gov/>). Design a kódování 2. verze programu připravil Christopher Gutteridge a dalšími spolupracovníky (Mike Jewell, Al Riddoch a Gui Power). Nejnovější verze programu (2.3.0), která přichází s dalším zlepšením řady funkcí, je právě testována. Instalace softwaru Eprints je relativně snadnou záležitostí, vyžaduje jistý prostor na webserveru, jistý čas na jeho instalaci ze strany správce systému a zejména pak jeho údržbu. Kvalitní budování archivu elektronických tisků vyžaduje, podle toho jak je veliký a jak je vnitřně strukturován, určitý tým editorů (ediční radu apod.), kteří, kromě jiného, zabezpečují také obsahové uspořádání archivu.

### **3 Archivy elektronických tisků v oboru informační věda**

V následující části budou uvedeny podstatné charakteristiky trojice existujících archivů elektronických tisků jako novodobých systémů pro saturování potřeb odborníků z oboru informační věda. Vznikly v nedávné době a jádro jejich systémů je provozováno na základě jediného aplikačního programu charakterizovaného v předcházející části. Texty jednotlivých dílčích částí prezentují podstatné znaky archivů. Kromě okolností jejich vzniku a provozu je největší pozornost věnována specifikům jednak vstupních procesů (typologii dokumentů, auto-archivaci a metadatovému popisu), jednak procesů výstupních (vyhledávacích a jiných služeb). V řadě parametrů jsou archivy mezi sebou porovnávány.

#### **3.1 Archiv elektronických tisků pro vědy o informacích a komunikaci „@rchiveSIC“**

Prvním ze tří archivů elektronických tisků je francouzský Archiv pro vědy o informacích a komunikaci @rchiveSIC (Archive en Sciences de l'Information et de la Communication, SIC, <http://archivesic.ccsd.cnrs.fr/>), který byl založen v jarních měsících roku 2002 a do úplného provozu uveden jako jeden z prvních v našem oboru již v červnu 2002. Jedním z podnětů k založení tohoto archivu byly alarmující výsledky vzešlé z ankety provedené francouzským ministerstvem výzkumu. Anketa sledovala využívání vědeckých či odborných periodik v oblasti věd o informacích a komunikaci (SIC). Bylo zjištěno, že ze 263 světových periodik pouze 17 titulů bylo známo alespoň jednomu odborníkovi ze dvou. Z 40 titulů přímo francouzských periodik jich bylo známo jenom 12 alespoň jednomu ze dvou odborníků [GALLEZOT, 2003, Des revues SIC].

Odpovědí na tuto situaci byl návrh na založení otevřeného archivu pro vědy o informacích a komunikaci. Jeho tvůrci a správci zároveň, tj. Gabriel Gallezot z Národní vysoké školy věd o informacích a knihovnách (ENSSIB) v Lyonu, Ghislaine Chartronová z Regionální jednotky pro výchovu v oblasti vědeckých a technických informací (URFIST) v Paříži a Jean-Max Noyer z Univerzity Paříž 7, využili příznivé situace vzniku Centra pro přímou vědeckou komunikaci (Centre pour la Communication Scientifique Directe, <http://ccsd.cnrs.fr>) v rámci francouzského Národního centra pro vědecký výzkum (CNRS), které již v té době provozovalo archiv pro elektronické disertace (<http://tel.ccsd.cnrs.fr>) a zrcadlové sídlo světového archivu pro přírodní vědy arXiv.org (<http://ccsd.cnrs.fr/arXiv.html>). Archiv @rchiveSIC byl založen na serveru výše zmíněného Centra [GALLEZOT, 2002, část VI]. Na jeho založení se dále významně podílely jednotky URFIST v Paříži a v Nice, Výzkumná skupina pro informační služby (GRESI), která do roku 2003 pracovala v rámci ENSSIB, a Výzkumné centrum pro komunikaci obrazových informací Oddělení pro vzdělávání v oblasti kinematografie, komunikace a informací (CCI/CRECI) na Univerzitě Paříž 7.

Francouzský archiv byl jedním z prvních, který v oboru informační věda adaptoval software Eprints. Byl a stále ještě je provozován v jeho verzi 1.1.1. Ve srovnání s jinými archivy má však lepší design prezentace informací a výstupů. Jako jediný umožňuje navíc vyhledávání v úplných textech. Adaptaci softwaru pro potřeby @rchiveSIC připravil Christian Rossi. Nová verze 2.2 zatím nebyla implementována, takže systém nevyužívá některých nových funkcí (například prezentaci citací z uloženého dokumentu a jejich propojování na úplné texty dokumentů - srovnej s verzí 2.2 nasazenou u archivů DLIST a E-LIS). Nicméně, archiv je již plně připraven na integraci archivů pomocí sdílení metadat na základě protokolu OAI-PMH ve verzi 2.0 [LAGOZE, 2003]. Archiv je od 19. prosince 2002 zaregistrován v centrále iniciativy OAI (<http://www.openarchives.org/Register/BrowseSites.pl>) včetně báze URL (<http://archiveSIC.ccsd.cnrs.fr/perl/oai20>). Metadata archivu jsou tudíž dosažitelná pro návazné informační služby včetně internetových vyhledávacích služeb, jako je Google.

Archiv @rchiveSIC je primárně určen komunitě frankofonních učitelů a badatelů z oblastí věd o informacích a komunikaci, lze ho tedy označit za mezinárodní (občas bývá označován jako archiv národní povahy). Dominuje v něm v současné chvíli francouzština, ukládat dokumenty je ale možné také v jazycích dalších (anglickém, německém, italském nebo španělském) [GALLEZOT, 2003]. K uložení elektronického dokumentu je nutná, tak jako v dalších archivech, registrace autora, který poskytuje základní údaje o sobě a na základě nich obdrží ze strany systému ID a heslo pro vstup do modulu pro ukládání dat. V rámci ukládacího procesu autor vyplňuje metadata o dokumentu pomocí webového formuláře.

Archiv má **vědecký a výzkumný ráz**. Přijímány jsou, pokud jde o **typy dokumentů**, pouze články nebo zprávy, které mohou mít trojí publikační status: 1. dokumenty připravené v rámci aktuálně prováděného výzkumu, kdy si autor přeje rychle sdílet nové poznatky, 2. dokumenty předložené v dané chvíli k recenzování z důvodu publikování ve vědeckém časopise nebo na konferenci a 3. dokumenty připravené již k publikování. Typologií přijímaných dokumentů se francouzský archiv výrazně liší od dalších dvou analyzovaných archivů. Nejvíce dodržuje definici termínu „elektronický tisk“. Elektronické dokumenty jsou přijímány ve 4 formátech: RTF, DOC, PDF a HTML. Formáty PDF a HTML jsou generovány automaticky, pokud autor uloží dokument ve zbývajících dvou formátech. Tuto službu nabízí zatím jenom francouzský archiv. Definitivní uložení dokumentu podléhá rozhodnutí odborného moderátora archivu, který ho může případně i odmítnout. Jde o minimální filtraci dokumentů danou cíli archivu, nikoliv o posuzování vědecké hodnoty jejich obsahu [GALLEZOT, 2003]. Lze říci, že se tím ve francouzském archivu udržuje vysoký standard kvality příspěvků a zběžná analýza obsahů zahrnutých dokumentů to potvrzuje. Archiv v současné době obsahuje velké množství velmi kvalitních dokumentů známých osobností francouzského výzkumu v oblasti věd o informacích

a komunikaci. Potěšitelné a inspirativní je, že se v něm již v tuto chvíli objevuje základní jádro výstupů z výzkumů v oblasti věd o informacích podporovaných finančně státem.

**Popis dokumentů** odpovídá standardu Dublinského jádra. Přesná specifikace prvků není v rámci veřejně dostupné dokumentace systému k dispozici, proto nebude na tomto místě interpretována. Podobá se však specifikacím v dalších analyzovaných archivech (viz části 3.2-3.3). Příklad výstupního metadatového záznamu dokumentu z roku 2003 je na obr. č. 3, jeho komunikát ve struktuře XML a sémantice Dublinského jádra, který byl získán pomocí příkazu GetRecord OAI-PMH přes kvalitní testovací systém iniciativy OAI „Repository Explorer“ (<http://oai.dlib.vt.edu/cgi-bin/Explorer/oai2.0/testoai>), je prezentován na obr. č. 4.

Pokud jde o věcné zpracování, pak autor především zařídí svůj dokument do některé z 21 nabízených **předmětových (tematických) kategorií (oblastí)**, jejichž přehled je v originále a v překladu k dispozici na obr. č. 1 a 2. Povinně musí vybrat alespoň jednu kategorii, volitelně může vybrat i další dvě. Pokud si nevybere z této nabídky, zařídí dokument pod kategorií „Ostatní“ a navrhne moderátorovi novou kategorii, která může být do seznamu doplněna. Každá předmětová kategorie má svého vlastního moderátora. V případě, že autor nezařadil dokument správně, provede opravu moderátor.



**@rchiveSIC**

**Liste par domaine : (Racine)**

- [Bibliométrie, scientométrie](#) (13)
- [Cinéma, art, esthétique](#) (22)
- [Collectivités territoriales](#) (40)
- [Communication et information scientifique](#) (37)
- [Conflits, Stratégie, Veille](#) (12)
- [Documentation](#) (29)
- [Droit de l'information/communication](#) (1)
- [Economie, industries culturelles](#) (19)
- [Edition électronique](#) (42)
- [Education, formation](#) (14)
- [Espace public](#) (34)
- [Gestion des connaissances](#) (20)
- [Géopolitique](#) (8)
- [Histoire de l'information/communication](#) (22)
- [Hypertextes, hypermédia](#) (39)
- [Ingénierie des systèmes d'information](#) (27)
- [Muséologie](#) (10)
- [Médias de masse](#) (36)
- [Organisation et communication](#) (35)
- [Sociologie de l'information/communication](#) (80)
- [Théories information/communication](#) (53)
- [Autres](#) (1)

Il y a **308** documents répartis dans les domaines en **630** entrées.

En cas de problèmes : [adminsIC@ccsd.cnrs.fr](mailto:adminsIC@ccsd.cnrs.fr) Commentaires : [coordoSIC@ccsd.cnrs.fr](mailto:coordoSIC@ccsd.cnrs.fr)

Obr. č. 1: předmětové kategorie archivu @rchiveSIC v původním jazyce uvedené v rámci rejstříku pro vyhledávání [převzato 2004-01-15]

Seznam předmětových kategorií archivu @rchiveSIC odráží aktuální přístup francouzských odborníků k vymezení oboru, který u nás označujeme termínem „informační věda“. Na rozdíl od amerického pojetí (prezentovaného například i v rámci archivu DLIST), které rozšiřuje název oboru v současné době o oblast informačních technologií (Information Science and Technologies), je francouzské pojetí typické jednak tím, že jde o více věd o informacích, jednak tím, že je název rozšířen v tomto případě o slovo „komunikace“ (se zdůrazněním na komunikaci ve vědě). V předmětové kategorizaci @rchiveSIC se tak objevuje řada příbuzných oborů (muzeologie, studia masmédií, sociologie, ekonomika, umění, estetika aj.).

Obr. č. 1 a 2 odráží aktuální stav počtu uložených dokumentů i jejich rozdělení do předmětových kategorií z hlediska obsahu. K 15. lednu 2004 bylo v archivu uloženo celkem 308 dokumentů, jimž bylo dohromady přiděleno 630 předmětových kategorií (v průměru je jeden dokument opatřen dvěma kategoriemi). Nejvíce dokumentů je zařazeno pod kategorii

Sociologie informací a komunikace (80 dokumentů zahrnujících zejména výsledky ze sociologických výzkumů), dále pod kategorií Teorie informací a komunikace (53), Elektronické publikování (42), Územní správní úřady (40 dokumentů věnovaných zejména komunikaci informací v rámci státní a místní správy), Hypertext a hypermédia (39), Vědecká komunikace a informace (37) a Masmédia (36).

Bibliometrie, scientometrie (13)	Management znalostí (20)
Film, umění, estetika (22)	Geopolitika (8)
Územní správní úřady (40)	Historie informací / komunikace (22)
Vědecká komunikace a informace (37)	Hypertexty, hypermédia (39)
Konflikty, strategie a inteligence (12)	Inženýrství informačních systémů (27)
Dokumentace (vyhledávání informací) (29)	Muzeologie (10)
Právní aspekty informací a komunikace (1)	Masmédia (36)
Ekonomika, kulturní průmysl (19)	Organizace a komunikace (35)
Elektronické publikování (42)	Sociologie informací/komunikace (80)
Výchova, (elektronické vzdělávání), školení (14)	Teorie informací/komunikace (53)
Veřejný prostor (34)	Ostatní (1)

Obr. č. 2: předmětové kategorie archivu @rchiveSIC v českém překladu [stav k 15. 1. 2004]

Věcný popis dále zahrnuje **abstrakt** a **volně tvořená klíčová slova** ve francouzštině (pravděpodobně povinné prvky) a v angličtině (pravděpodobně nepovinné). Jak dokládá srovnání jednotlivých prvků z obr. č. 3 a 4, je komunikace záznamů přes protokol OAI-PMH (na základě nekvalifikovaného Dublinského jádra) v současné chvíli nastavena tak, že exportovaný záznam nezahrnuje některé formální prvky (formát aj.) a klíčová slova ve francouzštině. Předmětové kategorie jsou exportovány pouze v anglické verzi.

Page d'accueil  
Lister  
Rechercher  
S'enregistrer  
S'abonner  
Déposer  
Vos documents  
Aide  
A propos

[Écrire un commentaire](#)

**ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la Communication : rôle, fonctionnement et usage.**  
Gabriel Gallezot. Article, Archimag - Guide Pratique - Les publications électroniques. 09 septembre 2003. Publié ou en cours de publication.

**Formats disponibles :** [DOC](#)  
[Adobe PDF](#)

Dépôt multi-fichier en HTML

- [Voir le document](#)
- [Récupérer les fichiers](#)
- [Détail des fichiers](#)

**Résumé :**

La création des archives ouvertes est née en réaction au paradoxe de l'édition scientifique : alors que le chercheur offre ces résultats, la communauté scientifique trouve un certain nombre d'obstacles pour lire ces publications. Pourtant la majorité des chercheurs accèdent aujourd'hui à Internet et pourraient se saisir de cette opportunité pour partager plus facilement les résultats de leurs travaux. De plus, l'inflation des écrits et la nécessité d'une validation dans un temps plus court pour certains résultats (souvent expérimentaux) se heurtent aux limites du modèle traditionnel de publication.

**Mots-clés :** publication électronique, archive ouverte,

**Domaine :** [Edition électronique](#)

**Code identifiant :** sic\_00000600

**Déposé par :** [Gabriel Gallezot](#) le 18 septembre 2003 (15:09)

[Écrire un commentaire](#)

Obr. č. 3: výstupní metadatový záznam preprintu příspěvku z komerční publikace uloženého v bázi archivu @rchiveSIC včetně nabídky k napsání komentáře [převzato 2004-01-15]

Archiv @rchiveSIC je sice primárně určen pro frankofonní odborníky, rozhraní systému je ale připraveno i v jazyce anglickém (včetně seznamu předmětových kategorií) pro uživatele z jiných jazykových oblastí. Poskytuje, kromě komplexních služeb pro autory ukládající dokumenty, řadu dalších **standardních i nadstavbových služeb**. Jde především o **vyhledávání informací**. Systém hned na úvodní stránce nabízí krátký seznam **aktuálně uložených dokumentů**, pokud byly uloženy. V rámci vyhledávání v režimu prohlížení nabízí zajímavý **rejstřík 10 nejvyužívanějších dokumentů** k danému datu a samozřejmě také rejstříky dokumentů řazených podle **autorů a předmětových kategorií**. V tomto archivu však, na rozdíl od ostatních, chybí rejstřík dokumentů upořádaných podle roku vytvoření/zveřejnění. **Přímé zadávání dotazu** je možné jak v jednoduchém režimu (podle slov z názvu a abstraktu ve francouzštině nebo klíčových slov ve francouzštině, názvu časopisu nebo konferenčního sborníku, autorů a předmětových kategorií dle výběru z řízeného seznamu), tak v režimu pokročilém (podle názvu dokumentu, autorů, názvu časopisu nebo konferenčního sborníku, slov z abstraktu ve francouzštině a angličtině a klíčových slov, předmětových kategorií dle výběru z řízeného seznamu a typu dokumentu). Specifikem francouzského archivu je vyhledávání uložených dokumentů podle identifikátoru (celého identifikátoru nebo jenom čísla). Je vhodné k rychlému nalezení záznamu a dokumentu, ale předpokládá se již dobrá orientace v archivu ze strany uživatele. Unikátní a velmi užitečné, opět ve srovnání s dalšími dvěma archivy, je **vyhledávání v úplném textu dokumentu**, připraveného na základě volně dostupného programu „ht://Dig“ (<http://www.htdig.org/>).

```
<record>
  <header>
    <identifier>oai:archiveSIC.ccsd.cnrs.fr:sic_00000600</identifier>
    <datestamp>2003-09-18</datestamp>
    <setSpec>sic_edit</setSpec>
  </header>
  <metadata>
    <oai_dc:dc
      xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
        http://www.openarchives.org/OAI/2.0/oai_dc/oai_dc.xsd">
      <dc:creator>Gallezot, Gabriel</dc:creator>
      <dc:description>La création des archives ouvertes est née en réaction au paradoxe de
        l'édition scientifique : alors que le chercheur offre ces résultats, la communauté
        scientifique trouve un certain nombre d'obstacles pour lire ces publications. Pourtant
        la majorité des chercheurs accèdent aujourd'hui à Internet et pourraient se saisir de
        cette opportunité pour partager plus facilement les résultats de leurs travaux. De
        plus, l'inflation des écrits et la nécessité d'une validation dans un temps plus court
        pour certains résultats (souvent expérimentaux) se heurtent aux limites du modèle
        traditionnel de publication. </dc:description>
      <dc:title>ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la
        Communication : rôle, fonctionnement et usage.</dc:title>
      <dc:language>fr</dc:language>
      <dc:subject>Electronic publishing</dc:subject>
      <dc:date>2003-09-09</dc:date>
      <dc:identifier>http://archiveSIC.ccsd.cnrs.fr/sic_00000600.en.html</dc:identifier>
      <dc:type>Text</dc:type>
    </oai_dc:dc>
  </metadata>
</record>
```

Obr. č. 4: základní část metadatového záznamu z archivu @rchiveSIC ve struktuře XML exportovaného dle protokolu OAI-PMH pomocí příkazu GetRecord [převzato 2003-12-19]

Pro registrované uživatele nabízí archiv **diskusní fórum**. Uživatelé mohou také zasílat komentáře k uloženým dokumentům (Écrire un commentaire - viz na obr. č. 3). Standardní službou je i poskytování **signálních služeb** (typu Alert) na základě zadaných profilů uživatelů



pomocí elektronické pošty s tím, že uživatel sice získává seznam metadatových záznamů, ale může si bezprostředně v prostředí WWW okamžitě stáhnout úplný text. Specifickou službou francouzského archivu je **konverze dokumentů** do formátu PDF a HTML ze strany pracovníků archivu, a to z důvodu jejich lepší viditelnosti a trvalého uložení (tuto službu neposkytují další dva analyzované archivy). Služba **poskytování metadatových záznamů** dalším vyhledávacím a jiným informačním službám na základě protokolu OAI-PMH již byla zmíněna výše v textu (viz exportovaný záznam na obr. č. 4).

Nezbytnou službou profesionálního charakteru je poskytování **statistik** uložených dokumentů a uživatelů archivu, kteří buď ukládají dokumenty, nebo pouze odebírají signální výstupy, a také využívání (či návštěvnosti) archivu, která má stoupající tendenci. Tvůrci a správci archivu se již nyní zamýšlejí nad otázkou, jak docílit většího nárůstu dokumentů - navrženo je například ukládání všech statí z kolokvií a konferencí a paralelní ukládání článků z odborných časopisů [GALLEZOT, 2003].

### **3.2 Archiv elektronických tisků pro informační vědu a technologie „DLIST“**

Druhým archivem elektronických tisků v oboru informační věda je americká Digitální knihovna informační vědy a technologií **DLIST** (Digital Library of Information Science and Technology, <http://dlist.sir.arizona.edu>), která byla uvedena do úplného provozu v říjnu 2002 (testování probíhalo již od 20. června 2002) ve spolupráci Školy informačních zdrojů a knihovní vědy a Arizonského centra pro zdravotnické vědy na Arizonské univerzitě (USA). Lze si všimnout, že v tomto případě je ve jménu archivu elektronických tisků uvedeno slovo „digitální knihovna“. Souvisí to zřejmě jak s rozšířeným záběrem přijímaných typů digitálních dokumentů, tak s budoucím plánovaným zaměřením celého systému vůbec [COLEMAN, 2003c]. K hlavním tvůrcům a správcům archivu patří Anita S. Colemanová a Paul Bracke z výše uvedených institucí. Základním cílem uvedených institucí bylo vytvoření otevřeného, přes WWW dostupného digitálního repozitáře pro uvedený obor. I když se ve jménu archivu vyskytuje pouze termín „informační věda“ (vedle termínu „technologie“), jeho tvůrci často v dokumentaci a literatuře zdůrazňují širší termín „knihovní a informační věda“ (Library and Information Science, LIS) [COLEMAN, 2003a].

Systém elektronického archivu je provozován v tomto případě již ve verzi 2.2 programového balíku Eprints.org, který je připraven na sklizeň metadat ze strany návazných služeb na základě protokolu OAI-PMH (bázové URL: (<http://dlist.sir.arizona.edu/perl/oai2>)). K 15. 1. 2004 ale nebyla ještě provedena registrace archivu v centrále OAI. Systém běží pod operačním systémem Linux. Databáze je uložena v systému MySQL a prezentace informací na WWW je zajišťována pomocí webserveru Apache a programovacího jazyka Perl. Doplňkový program Webalizer (otevřený zdroj) zajišťuje analýzy a zprávy o využívání archivu ze strany uživatelů.

**Typologie zahrnovaných materiálů** (dokumentů) je ve srovnání s francouzským archivem velmi široká. Zjevně nejde jenom o elektronické tisky. Zahrnuje různé typy digitálních materiálů, ale jeho specifikem je větší množství typů materiálů určených pro potřeby **výuky a informační gramotnosti**. Do archivu mohou být zařazovány bibliografické instruktážní materiály (bibliografie a přehledové seznamy), knihy a jejich kapitoly, edice monografických materiálů, konferenční sborníky, konferenční referáty, konferenční vývěsky (postery), datové soubory (například z bibliometrických výzkumů), zprávy fakult nebo ústavů ze sledovaných oborů, průvodce, interaktivní materiály, časopisy (online stránkované i nestránkované) a jejich články, knihovnické instruktážní materiály (včetně materiálů typu „help“), články z novin a magazínů, preprinty, prezentace, technické a výzkumné zprávy, disertace, sylaby, osnovy kursů a plány cvičebních lekcí, výukové materiály (k počítačovým programům a k

tvorbě a využívání databází) a jiné. Pokud jde o publikační status ukládaných materiálů, může jít o materiály již publikované (nutné je v tomto případě dodání dokumentu v digitální formě vytvořené autorem, nikoliv vydavatelem, např. u souborů PDF), materiály, které mají být publikovány, nebo i materiály nepublikované.

Archiv DLIST zahrnuje v současné době dokumenty v angličtině [COLEMAN, 2003a, část 1]. Autor ukládá elektronický dokument přímou cestou na server (uploading), popřípadě může specifikovat URL, z něhož může být dokument transferován pracovníky archivu. V obou případech bude kopie díla uložena v databázi systému DLIST. Autor zároveň vyplňuje metadata, jejichž současná specifikace vychází z Dublinského jádra. Zahrnuje (prvky uvedené kurzivou nejsou povinné):

1. Typ elektronického tisku (viz položky typologie materiálů)
2. *Novější verze k (identifikátor předchozí verze dokumentu uloženého v archivu)*
3. *Komentář k (identifikátor komentovaného dokumentu uloženého v archivu)*
4. Status (nepublikován, v tisku, publikován)
5. *Veřejná doména (vyplňuje se v případě, že jde o starší a až nyní zveřejňovaný dokument jiného autora)*
6. *Jména editorů*
7. Název (dokumentu uloženého v archivu)
8. Předmětové kategorie (povinně jedna kategorie, maximálně dvě navíc volitelně)
9. Rok (zveřejnění nebo výroby ukládaného dokumentu)
10. *Abstrakt*
11. Konference (úplné jméno konference)
12. *Edice (název edice nebo souboru knih)*
13. *Datum konference*
14. *Lokace (místo konání konference)*
15. *Ročník (časopisu, v němž byl zveřejněn článek)*
16. *Číslo (časopisu, v němž byl zveřejněn článek)*
17. *Nakladatel*
18. *Komentář k (úplný bibliografický záznam komentovaného dokumentu uloženého v jiném archivu)*
19. *Alternativní lokace (URL jiných míst ukládaného dokumentu)*
20. *Klíčová slova (volně tvořená)*
21. *Doplňkové informace*
22. *Bibliografické odkazy (seznam citované/odkazované literatury určený pro potřeby tvorby citačního rejstříku)*
23. *Poznámky a návrhy (určené editorům archivu)*
24. *Formát (PDF, HTM, HTML, XLS, PPT, TEXT, Rich Text Format, PostScript, ASCII, DOC)*
25. *Doplňkové informace k formátu uloženého dokumentu*
26. *Stupeň bezpečnosti (lze limitovat přístup k dokumentu pouze pro personál archivu)*

Po vyplnění metadat připojuje autor elektronický soubor, popř. více souborů, a ukončuje celou vstupní proceduru odesláním příslušné funkce směrem do systému. Autor tím také předává práva na uložení svého dokumentu v archivu a na jeho volné využívání ze strany uživatelů. Je však nezbytné, aby si autor sám hlídal případnou kolizi práv ke zveřejnění dokumentu, zejména pokud jde o komerční vydavatele jeho dokumentu.

Obsah repozitáře archivu DLIST tvoří v tuto chvíli digitální dokumenty z oblasti knihovnictví a informační vědy (Library and Information Science, LIS) a informačních technologií (Information Technologies, IT). Věcný popis zahrnuje především zařizování dokumentu do alespoň jedné z nabízených předmětových kategorií. **Taxonomie** archivu DLIST, prezentovaná v rámci prohlížení rejstříku předmětových kategorií, zahrnuje v současné době 83 navržených položek, z nichž se po ročním provozu uplatnilo 48 (viz obr. č. 5). Seznam předmětových kategorií je otevřený a ukládající autoři mohou případně navrhnout položky nové. Nicméně, již na počátku tvůrci archivu vymezili dvě základní oblasti, kterým bude

věnována mimořádná pozornost: 1. **informační (a počítačová) gramotnost**, (jde zejména o zprávy, bibliografické informační materiály, instruktážní příručky apod.), 2. **informetrie (a bibliometrie)** (jde zejména o publikace z výzkumů a doprovodné datové soubory). Aktuální přehled četností obsazených kategorií prozrazuje obr. č. 5 (k 15. lednu 2004). V archivu se k tomuto datu nacházelo celkem 97 dokumentů, kterým bylo přiděleno dohromady 179 kategorií (v průměru šlo o necelé 2 kategorie přidělené jednomu dokumentu). Roční nárůst dokumentů se v roce 2003 (27 dokumentů) oproti roku 2002 (29 dokumentů) prakticky nezměnil a jde o nárůst velmi malý ve srovnání s dalšími dvěma archivy. Příčin, proč je americký archiv tak málo v pozornosti většího počtu autorů z oblasti informační věda, může být více. Jednou z nich může být i fakt, že obor je v USA zabezpečen velkým počtem komerčních produktů (včetně časopisů) a služeb, které jsou jim snadno dostupné. Autoři tudíž nejsou, alespoň prozatím, motivováni k ukládání kopií dokumentů, pokud by tomu i vydavatel nebránil, do volně dostupného archivu. Prozatím se v archivu DLIST nalézá největší počet dokumentů věnovaných digitálním knihovnám (16), informační gramotnosti ve spojení s knihovnickými instrukcemi (dohromady 16), bibliometrii ve spojení s informetrikou (14), studiím v oblasti vědy a technologií (13) a problematice akademického knihovnictví (8).


Akademické knihovny (8)	Hypertext a hypermédia	Mediální studia
Antropologie	Indexace (1)	Lékařské knihovny (2)
Archeologie	Informační analýza	Metadata (5)
Archivy (2)	Extrakce informací	Muzea
Umělá inteligence (1)	Informační gramotnost (9)	Zpracování přirozeného jazyka
Bibliometrie (11)	Informační věda (6)	Neurověda
Katalogizace (6)	Chování při hledání informací (4)	Ontologie
Citační analýza (2)	Informační systémy (5)	Filozofie
Třídění (3)	Informetrie (3)	Psychologie
Kocitační analýza (2)	Mezioborovost (2)	Kvalitativní výzkum
Kognitivní věda	Internet (3)	Kvantitativní výzkum
Komunikace (1)	Zurnalistika	Referenční služby
Počítačová lingvistika	Management znalostí	Výzkumné metody (1)
Počítačová věda (1)	Organizace znalostí (2)	Vědecká komunikace (3)
Dolování znalostí (1)	Reprezentace znalostí (3)	Studia ve vědě a technologiích (13)
Instrukce pro vyhledávání v databázi	Struktury znalostí (3)	Sociální epistemologie
Databáze (5)	Věda o učení (4)	Sociologie (1)
Digitální knihovny (16)	Knihovny (6)	Software (5)
Distribuované učení (3)	Výchova v knihovní a informační vědě (2)	Standardy
Ekonomika (1)	Knihovnické instrukce (7)	Školení
Informační ekonomie (2)	Knihovní věda (2)	Studia uživatelů (4)
Elektronické publikování (6)	Knihovnická statistika	Virtuální komunity
Epistemologie	Knihovnické systémy (4)	Měření webu (1)
Evaluace	Lingvistika	Dolování webu (2)
Geografické digitální knihovny	Management (1)	Bezdrátové technologie
Věda o geografických informacích	Informační systémy v managementu	World Wide Web (3)
Geografie	Knihovnictví v oblasti map (1)	XML
Komunikace člověk-počítač (4)		Z39.50

Obr. č. 5: Taxonomie archivu DLIST v češtině se zachováním původního pořadí podle anglické abecedy [převzato 2004-01-15]

Archiv DLIST nabízí řadu standardních **služeb**, nabídka ovšem zatím není tak bohatá jako u archivu francouzského. Služby pro ukládající autory byly již zmíněny výše. V rámci **vyhledávání informací** nabízí archiv prioritně prohlížení seznamu **nových přírůstků** za poslední týden. Dále lze vyhledávat informace pomocí rejstříků dokumentů uspořádaných podle roku zveřejnění a podle předmětové kategorie. Předmětovou kategorii si uživatel vybírá klikacím způsobem, smysl má ovšem jenom klikání na položky, které mají uveden alespoň jeden výskyt v databázi indikovaný příslušnou arabskou číslicí. Po získání úplného záznamu (viz obr. č. 6) lze v dalších krocích buď volit úplný text dokumentu, nebo navigovat dále pomocí hyperodkazů na další záznamy (minimálně podle předmětové kategorie). Protože

archiv DLIST je provozován již ve verzi 2.2, je možné již u některých záznamů vidět také navigaci ze záznamů citované literatury pomocí odkazu SEEK, stejně jako tomu je i u archivu E-LIS (viz další dílčí část). **Přímé formulování dotazu** lze realizovat buď v jednoduchém režimu (podle slov z názvu, abstraktu a klíčových slov, autorů a editorů a roku zveřejnění) nebo v režimu pokročilém (podle slov z názvu dokumentu, autorů, slov z abstraktu, klíčových slov, předmětových kategorií, typu dokumentu, jména konference, pořadatelské organizace konference, editorů, publikačního statusu, vydavatele a roku zveřejnění).

Registrovaní uživatelé mohou využívat jednak **diskusního listu** archivu (DLIST listserv), jednak mohou získávat **signální služby** (typu Alert) na základě zadaných profilů pomocí elektronické pošty.



[Home](#) | [About](#) | [Browse](#) | [Search](#) | [Register](#) | [User Area](#) | [Credits](#) | [Help](#)

---

## Presenting DLIST

Coleman, Anita (2003) Presenting DLIST.

Full text available as:  
[Microsoft Powerpoint Presentation \(.ppt\)](#) - Archive staff only - Requires Microsoft Powerpoint

### Abstract

Forthcoming

EPrint Type: Presentation

Subjects: [Academic Libraries](#)

ID Code: 184

Deposited On: 19 March 2003

*Obr. č. 6: výstupní minimální metadatový záznam prezentace A. Colemanové ve formátu PPT uložené do repozitáře archivu DLIST dne 19. března 2003; přístup k dokumentu je (dočasně) omezen pouze pro personál obsluhující archiv [převzato 2004-01-15]*

**Výstupy statistického charakteru** patří k velmi sledovaným jak ze strany archivu, tak ze strany archivů konkurenčních, ale i ze strany běžných koncových uživatelů. Statistika využívání archivu DLIST od jeho počátku k lednu 2003 vykazala, kromě jiného, že 6. místo ve využívání obsadili uživatelé z České republiky (po USA a Kanadě) [COLEMAN, 2003a, část 8].

Jak bylo uvedeno výše, není archiv DLIST prozatím příliš vyhledávaným archivem ze strany autorů dokumentů. Jeho celkový fond je nejmenším ze všech tří analyzovaných archivů. I to byl zřejmě jeden z důvodů jeho nové plánované inovace, která míří k rozšíření archivu o nové moduly, které by dohromady vytvořily nový systém vyššího řádu nazvaný „**DLISTConnection**“. Koncepte plánovaného systému je představena v přípravné studii ke grantovému projektu, která je uložena v archivu DLIST [COLEMAN, 2003c]. Systém DLISTConnection je navrhován jako budoucí **služba pro informační a technologickou gramotnost** (ITL, Information and Technology Literacy), která by měla významně podpořit gramotnost ve vědě a v oblasti zdravotnictví a která by se zároveň stala významnou komponentou systému americké Národní vědecké digitální knihovny NSDL (National Science Digital Library, <http://nsdl.org/>). Autoři projektu předpokládají, že k poskytování komplexních služeb systému DLISTConnection bude vytvořen jako jádro speciální digitální

fond ITL materiálů, tj. databáze cca 3000 výukových objektů (instruktážní materiály, dokumentace k softwarům, průvodce a manuály k různým databázím a jiné výukové objekty). Předpokládá se jejich precizní metadatový popis. Jak konstatují autoři projektu, právě tyto typy materiálů nejsou v současné době předmětem systematického sběru systému NSDL [COLEMAN, 2003c, I. Purpose]. Pro splnění tohoto cíle bude nezbytné stávající specifikaci metadat archivu DLIST rozšířit o další prvky. Plánována je také podrobná specifikace metadat pro potřeby autorskoprávní ochrany získaných materiálů (řada z nich bude transferována do systému na základě dohod s komerčními producenty těchto materiálů). Nezbytnou součástí projektu bude také příprava automatického toku všech informací včetně metadat na základě principů a protokolů OAI. Testována bude i interoperabilita archivů provozovaných v různých aplikačních programech (jde zejména o programy Eprints.org a DSpace). K testování výstupů systému DLISTConnection budou vybrány speciální komunity americké populace - vedle profesionálů působících v oblasti informační a technologické gramotnosti půjde zejména o reprezentanty komunity afrických Američanů a původních Američanů, u kterých se předpokládá velmi citlivé a tudíž účinné vyhodnocování jednotlivých výstupů systému DLISTConnection [COLEMAN, 2003c, VIII. Impact and Expected Results].

### **3.3 Archiv Elektronické tisky v knihovní a informační vědě „E-LIS“**

Archiv Elektronické tisky v knihovní a informační vědě **E-LIS** (Eprints in Library and Information Science, <http://eprints.rclis.org>) je nejmladším z trojice analyzovaných otevřených archivů v oboru knihovní a informační věda. Do provozu byl sice uveden až v lednu 2003, podařilo se mu však k dnešnímu datu získat první místo co do objemu uložených dokumentů. Archiv E-LIS byl založen jako rozšíření již existující služby RCLIS (Research in Computing, Library and Information Science, <http://rclis.org/>) a zejména jeho dílčí služby DoIS (<http://dois.mimas.ac.uk/>), které bude věnována následující 4. část textu. Cílem bylo podpořit otevřený přístup k dokumentům reprezentujícím výzkumné aktivity v oboru informační věda v globálním rámci. Má otevřený mezinárodní ráz, autoři v něm mohou publikovat dokumenty bez geografických nebo jazykových omezení [BARRUECO, 2003b, část 4]. Základním doporučeným jazykem je nicméně angličtina. Pokud je dokument v jiném jazyce, požaduje se uložení minimálně abstraktu v angličtině. Z hlediska obsahu prozatímni množiny uložených dokumentů by archiv E-LIS mohl být označen i jako archiv evropský, velké množství literatury je například italské nebo španělské provenience, uloženy byly také již dokumenty v srbochorvatštině. Systém zahrnuje jak publikovanou tak nepublikovanou literaturu z oboru knihovní a informační věda (včetně informačních technologií a příbuzných disciplín). V tom se podobá archivu DLIST.

Archiv E-LIS je provozován v nejnovější revidované verzi 2.2.1 programu Eprints.org. Je rovněž plně připraven na integraci archivů pomocí sdílení metadat na základě protokolu OAI-PMH (verze 2.0). Od 20. ledna 2003 je zaregistrován v centrále iniciativy OAI včetně báze URL (<http://eprints.rclis.org/perl/oai2>). Metadata archivu jsou tedy dosažitelná pro všechny návazné informační služby. Na automatický sběr metadat z tohoto archivu je rovněž připravována služba DoIS (viz další část). Systém je provozován na serveru italské meziuniverzitní organizace CILEA. Využit je operační systém Linux (RedHat 7.3). Databáze je uložena v databázovém systému MySQL a prezentace informací na WWW je zajišťována dynamicky pomocí web serveru Apache a souboru skriptů v jazyce Perl.

Archiv E-LIS je otevřený, jak deklarují jeho představitelé, maximálnímu počtu autorů [BARRUECO, 2003b, část 4]. I v rámci tohoto archivu je od autorů požadována registrace před prvním uložením dokumentu. Příspěvky podléhají kontrole ze strany editorů z ediční rady systému. Přijímány jsou v podstatě jakékoliv dokumenty včetně sekundárních, jak dokládá i dále uvedená **typologie dokumentů**. Podobá se téměř zcela typologii archivu

DLIST, to znamená, že rovněž v jistém smyslu překračuje definici termínu „elektronický tisk“. Uloženy tedy mohou být bibliografie, knihy a jejich kapitoly, edice monografických materiálů, konferenční sborníky, konferenční referáty, konferenční vývěsky (postery), datové soubory, zprávy fakult nebo ústavů ze sledovaných oborů, průvodce, časopisy (online stránkované i nestránkované) a jejich články, knihovnické instruktážní materiály, články z novin a magazínů, preprinty, prezentace, technické a výzkumné zprávy, sylaby, disertace, výukové příručky a jiné. K typu průvodce jsou zahrnovány ještě příručky (manuály). Zahrnuty nejsou, oproti archivu DLIST, interaktivní materiály a přehledové seznamy. Navíc se vyskytují projektové a obchodní plány. Rovněž typy publikačních statusů jsou stejné: materiály již publikované, materiály, které mají být publikovány, nebo materiály nepublikované. Archiv je tedy vůči autorům (velké množství z nich tvoří knihovníci), dá se říci, velice vstřícný a tak se poměrně rychle zaplňuje příspěvky. Na počátku roku 2004 jich bylo již více než 400. V tuto chvíli lze konstatovat, že jeho obsah odráží velmi čilý aktuální dění konané v daném oboru (určitě) v evropském kontextu. V řadě případů jde i o kratší popisné příspěvky nebo PPT prezentace z konferencí, které primární výzkum a vývoj odrážejí jenom sekundárně (v několika případech se vyskytují duplicity jednoho základního příspěvku). V tuto chvíli to snad není na závadu, ovšem pokud by se archiv rychle rozrůstal, není vyloučeno, že vyvstane potřeba nastavení dalších limitů pro příjem dokumentů.

Archiv přijímá dokumenty ve velkém množství formátů: PDF, PostScript, TeX, LaTeX (DVI), HTML, XML, ASCII, PPT, DOC a RTF, avšak alespoň jeden z formátů dokumentu, který je ukládán, musí být ve formátu HTML, PDF, PostScript nebo ASCII (důvodem je extrakce záznamů citací z dokumentů automatickou cestou). Doporučovanými formáty jsou však formáty PDF a HTML.

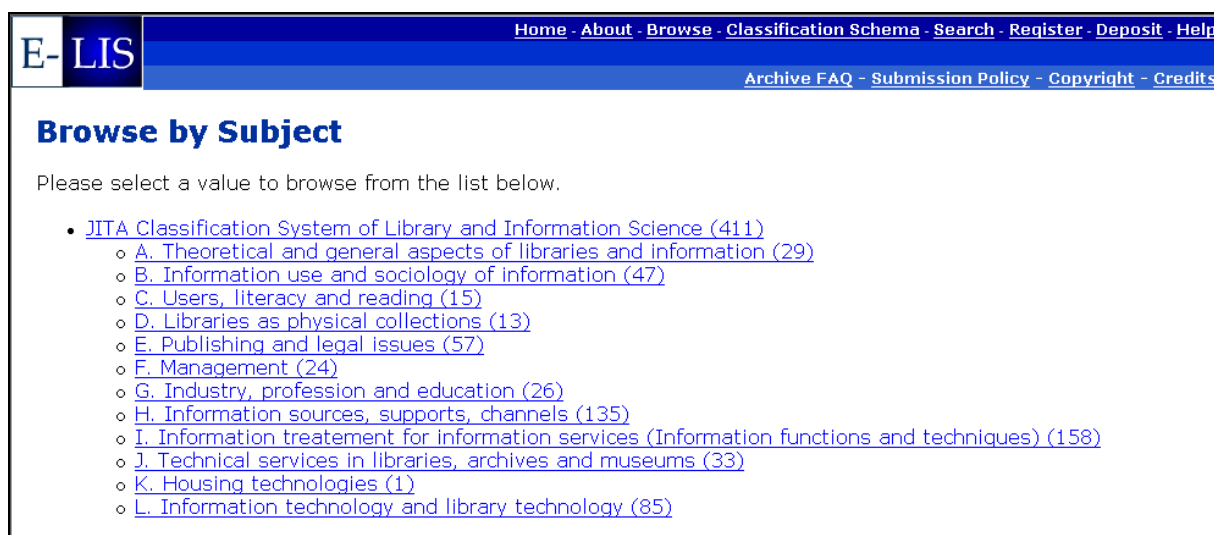
Před vlastním uložením dokumentu autor vyplňuje **metadata** o dokumentu. Jejich skladba je podobná jako v archivu DLIST, takže výčet prvků na tomto místě není nutné opakovat.

Otázky **copyrightu** jsou ze strany archivu neustále sledovány. Dokumenty uložené v archivu jsou chráněny na základě copyrightu, který drží autor sám (v případě preprintu). V případě postprintů se autoři sami musejí postarat o to, aby ukládaný dokument nebyl nějakým způsobem omezen z hlediska copyrightu. Jde zejména o dokumenty publikované dříve někde jinde (pak záleží na formulaci podepsané smlouvy s vydavatelem a na jeho případném dodatečném souhlasu s uložením dokumentu do otevřeného archivu elektronických tisků). Je doporučováno, aby byl do archivu například uložen příspěvek ve vlastním PDF formátu, nikoliv formátu, který připravil nakladatel časopisu, v němž byl příspěvek publikován.

Specifikem archivu E-LIS je jeho **třídění JITA** (<http://rclis.org/internal/jita.txt>). Jde v podstatě zatím jen o jednoduché schéma předmětových kategorií. Tvůrci systému zatím kvůli jednoduchosti navrhli jenom jedinou úroveň dělení. Třídění bylo připraveno na základě sloučení původního třídění „NewAgentTopic“, které vytvořil a do roku 1998 udržoval Mike Keen (Aberystwyth, Velká Británie), a třídění německého časopisu „Review of Information Science (RIS)“, které původně vytvořil Donald Soergel. Zkratka JITA pochází z křestních jmen zakladatelů archivu E-LIS i jeho třídění (José Manuel Barrueco Cruz, Imma Subirats Coll, Thomas Krichel a Antonella De Robbio).

Ve srovnání s výše charakterizovanými dvěma archivy je třídění JITA více propracované. Zahrnuje v současné době 12 hlavních předmětových kategorií, označených písmeny velké abecedy (originál schématu viz obr. č. 7) a doplněných slovní hodnotou. Kategorie jsou v celkovém schématu zařazeny do tří úrovní (viz obr. č. 8). Kategorie A-B patří do obecné úrovně (zahrnuje teoretické otázky a obecnosti), kategorie C-G do střední úrovně (zahrnuje otázky uživatelů informací, typů knihoven, otázky publikování a právních aspektů, řízení, informačního průmyslu, profese a výchovy v oboru) a kategorie H-L do úrovně specifické

(zahrnuje konkrétní otázky informačních zdrojů, zpracování informací, informačních a technických služeb v knihovnách, archivech i muzeích a technologií). Základní schéma, publikované na stránkách sídla archivu, je vnitřně doplněno výčtem dílčích předmětových kategorií (celkem 124) zahrnutých do kategorií základních. V průměru tedy na jednu základní kategorii připadá asi 10 dílčích (nejsou ale nijak dále pořádány). Jejich úlohou je zřejmě jenom orientace pro autory při výběru konkrétní základní kategorie (popř. dalších 2) pro jimi popisovaný a ukládaný dokument. Při bližším náhledu lze zjistit jisté duplicity či překrývání dílčích kategorií napříč základními kategoriemi, což by mohlo uživatele někdy zmást. Nicméně celé schéma je, jak sami tvůrci sdělují, otevřené a jeho úpravu a zdokonalování lze předpokládat [BARRUECO, 2003b, část 4.4]. O schématu je možné diskutovat, nicméně je třeba ocenit úsilí tvůrců elektronického archivu při pokusu o nové oborové pořádání informací. Je dokladem toho, že ani novodobé systémy se bez profesionálního pořádání nemohou obejít, ačkoliv i na to nejsou jednotné názory. V již dříve citované dokumentaci softwaru Eprints.org jsou zájemci ujišťováni, že v archivech elektronických tisků není klasifikací zapotřebí, neboť je nepravděpodobné, že by klasifikační systémy, které známe z významných oborových databází, přinesly lepší výsledky, než je booleovské vyhledávání dokumentů přes index plných textů googlovského stylu, zejména pokud je obohaceno o novodobé výpočty citovanosti, obsahové relevance, využívání ze strany uživatelů apod. [Eprints.org, 2002, část 26, Classification].



Obr. č. 7: základní schéma třídění JITA prezentované v rámci předmětového rejstříku archivu E-LIS [2004-01-15]

K 15. lednu 2004 archiv obsahoval již 411 dokumentů. Z hlediska obsahového jim bylo přiděleno celkem 623 předmětových kategorií (v průměru 1,5 kategorie na jeden dokument). Aktuální obsazení předmětových kategorií prozrazuje, že největší množství uložených dokumentů náleží do tříd specifické úrovně. Zahrnují zejména praktické a technologické otázky rozvoje informačních zdrojů, jejich systémů a služeb. Největší obsazení má kategorie Zpracování informací pro informační služby (Informační funkce a techniky) (158 dokumentů), dále kategorie Informační zdroje, nosiče a kanály (135 dokumentů), kategorie Informační a knihovnické technologie (85), kategorie Publikování a legislativní otázky (57) a kategorie Využívání informací a sociologie informací (47). Kategorie obecné i střední úrovně jsou obsazovány výrazně méně. Obsazování prakticky a technologicky orientovaných kategorií pravděpodobně souvisí i s výše uvedenou skutečností týkající se vstřícného přístupu archivu k autorům ukládaných dokumentů.

[Teoretické a obecné aspekty]: obecná úroveň	
A.	Teoretické a obecné aspekty knihoven a informací (29)
B.	Využívání informací a sociologie informací (47)
[Uživatelsky orientované aspekty, řídicí a manažerské funkčnosti]: střední úroveň (včetně sociálně ekonomických a právních otázek)	
C.	Uživatelé, gramotnost a četba (15)
D.	Knihovny jako fyzické fondy (13)
E.	Publikování a legislativní otázky (57)
F.	Management (24)
G.	Průmysl, profese a vzdělávání (26)
[Objekty, praktické a technologické aspekty]: specifická úroveň	
H.	Informační zdroje, nosiče, kanály (135)
I.	Zpracování informací pro informační služby (Informační funkce a techniky) (158)
J.	Technické služby v knihovnách, archivech a muzeích (33)
K.	Technologie budov (1)
L.	Informační a knihovnické technologie (85)

Obr. č. 8: základní schéma třídění JITA archivu E-LIS v českém jazyce [2004-01-15]

Roční nárůst dokumentů se v roce 2003 (157 dokumentů) oproti roku 2002 (90 dokumentů) výrazně změnil, jde o nárůst relativně vysoký ve srovnání s dalšími archivy (poznamenat je však třeba, že v daném čase může do archivu přibýt i starší dokument, takže čísla uváděná v rámci rejstříku roků nevypovídají přesně o ročním nárůstu dokumentů do archivu).

E-LIS

[Home](#) - [About](#) - [Browse](#) - [Classification Schema](#) - [Search](#) - [Register](#) - [Deposit](#) - [Help](#)  
[Archive FAQ](#) - [Submission Policy](#) - [Copyright](#) - [Credits](#)

## A Pioneering Spirit: Using Administrative Metadata to Manage Electronic Resources

Medeiros, Norm (2003) A Pioneering Spirit: Using Administrative Metadata to Manage Electronic Resources. *OCLC Systems & Services* 19(3).

Full text available as:  
[PDF](#) - Requires [Adobe Acrobat Reader](#) or other PDF viewer.

### Abstract

This article describes administrative metadata, and its use in managing electronic resources. The focus of the article is an interview with Tim Jewell, Head of Collection Management Services at the University of Washington and Adam Chandler, Information Technology Librarian at Cornell University.

**Keywords:** Electronic Resource Management Initiative, ERMI, administrative metadata; electronic resources; e-resources; Tim Jewell, Adam Chandler

**Subjects:** [L. Information technology and library technology](#)  
[J. Technical services in libraries, archives and museums](#)

**ID Code:** 400

**Deposited By:** [Medeiros, Norm](#)

**Deposited On:** 27 October 2003

Anderson, R. & Zink, S.D. (2003). "Implementing the unthinkable: the demise of periodical check-in at the University of Nevada." *Library Collections, Acquisitions, and Technical Services*, vol. 27, no. 1, p. 61-71. [SEEK](#)

Gorman, M. (1991). "The academic library in the year 2001: dream or nightmare or something in between?" *Journal of Academic Librarianship*, vol. 17, no. 1, p. 4-9. [SEEK](#)

Schulz, N. (2001). "E-journal databases: a long-term solution?" *Library Collections, Acquisitions, and Technical Services*, vol. 25, no. 4, p. 449-459. [SEEK](#)

Obr. č. 9: výstupní metadatový záznam preprintu článku z časopisu uloženého v archivu E-LIS s odkazy na úplný text a záznamy citované literatury [převzato 2004-01-15]



Archiv E-LIS nabízí rovněž řadu standardních **služeb** pro koncové uživatele. V rámci **vyhledávání informací** nabízí prioritně již z úvodní stránky prohlížení seznamu **nových přírůstků** za poslední týden. Dále lze vyhledávat informace pomocí rejstříků dokumentů uspořádaných podle roku zveřejnění a podle předmětové kategorie. Po získání úplného záznamu (viz obr. č. 9) lze volit úplný text dokumentu nebo navigovat dále pomocí hyperodkazů na další záznamy (podle autora, předmětové kategorie) a také (u některých záznamů) pomocí odkazů u citované literatury (viz odkaz SEEK na obr. č. 9). **Přímé formulování dotazu** lze realizovat buď v jednoduchém režimu, a to buď přímo z hlavní stránky (podle slov z názvu, abstraktu a klíčových slov), nebo ze stránky pro jednoduché vyhledávání (podle slov z názvu, abstraktu a klíčových slov, autorů a editorů a roku zveřejnění), nebo v režimu pokročilém (podle slov z názvu dokumentu, autorů, slov z abstraktu, klíčových slov, předmětových kategorií, typu dokumentu, jména konference, pořadatelské organizace konference, editorů, publikačního statusu, vydavatele a roku zveřejnění).

Registrovaní uživatelé mohou využívat **diskusního listu** archivu a mohou rovněž získávat **signální služby** (typu Alert) pomocí elektronické pošty na základě zadaných profilů.

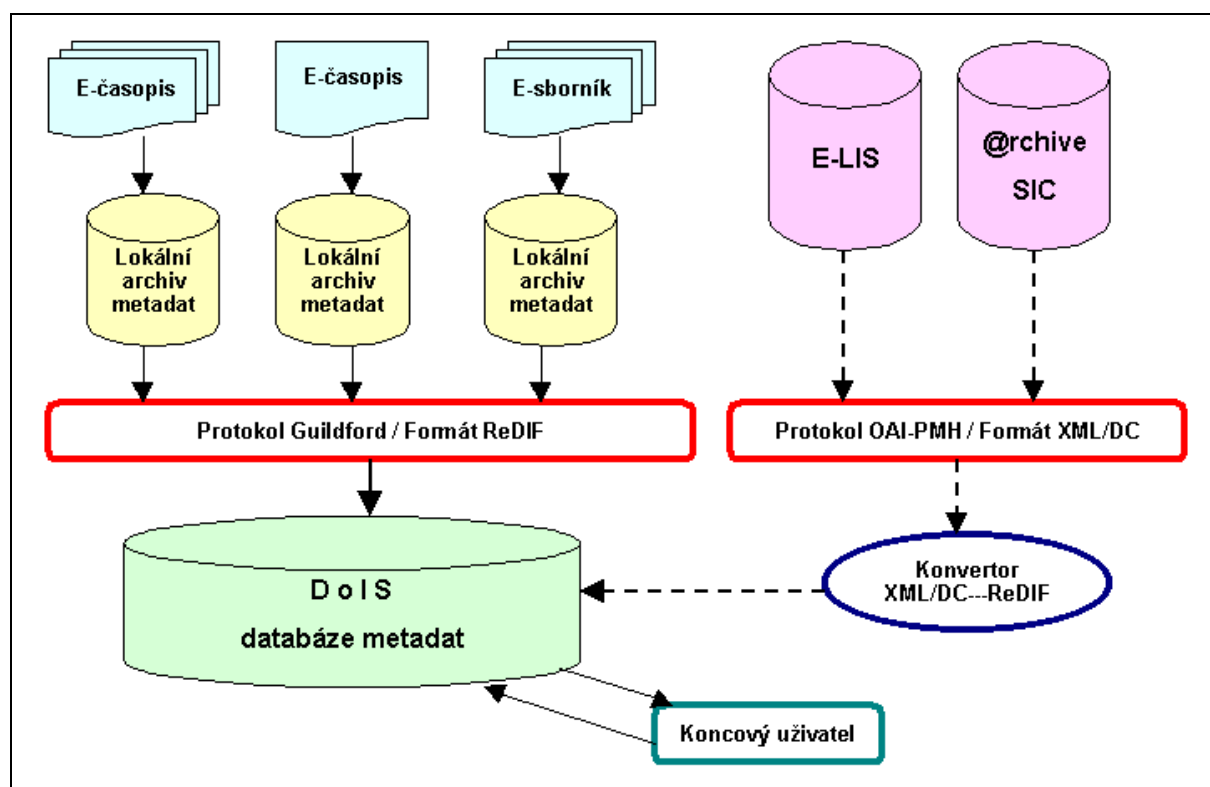
Výše zmíněná služba **propojování citovaných dokumentů** s jejich plnými texty (viz odkaz SEEK na obr. č. 9) je novinkou druhé verze programu Eprints.org. Služba je realizována na základě specifického programu „**ParaCite Toolkit**“ (zkráceně „**ParaTools**“, <http://paracite.eprints.org>), který je v současné chvíli integrální součástí programu Eprints [De ROBBIO, 2003, Reference Linking]. Program je schopen jednak extrahovat záznamy citací z archivovaných dokumentů (podmínkou je, že jsou ve formátech PDF, PostScript nebo HTML), jednak rozkládat záznamy citací na jednotlivá metadata a směřovat je ke službě s webovým formulářem pro vyhledávání příslušných dokumentů ve vybraných volně dostupných zdrojích na WWW. Patří k nim báze CiteBase (<http://citebase.eprints.org/cgi-bin/search>), báze systému ResearchIndex (<http://citeseer.nj.nec.com/cs>) a index vyhledávače Google (<http://www.google.com>). Do hry mohou v budoucnu vstoupit zdroje další. Celá procedura má stále experimentální charakter (ne vždy je uživatel uspokojen). Testuje se také propojování na bázi připravovaného standardu OpenURL.

#### **4 Elektronická knihovna a alternativní bibliografická služba pro informační vědu „DoIS“**

Pod zkratkou „**DoIS**“ se skrývá jméno specifické elektronické knihovny hybridního charakteru pro informace o dokumentech z oboru informační věda a částečně i pro dokumenty samotné, jsou-li k dispozici volně na WWW na mateřských serverech. Je třeba zdůraznit, že jde především o dokumenty pocházející z vědeckovýzkumné činnosti. Plné jméno systému zní „**Dokumenty v informační vědě**“ (Documents in Information Science). O hybridní charakter jde proto, že systém zahrnuje nejen záznamy dokumentů, které jsou volně dostupné v prostoru Internet, ale také záznamy dokumentů, které dostupné volně nejsou (jde zejména o analytické dokumenty z komerčně dostupných časopisů nebo sborníků). Systém bývá označován také termínem „gateway“ nebo „portál“, zejména pak ale termínem „**služba**“. O charakter služby v tomto případě půjde nejvíce a budeme mít na mysli především služby bibliografického charakteru (ve formě komunikace metadat o dokumentech). Systém byl založen a do provozu uveden již v roce 1999, a to, jak bylo uvedeno v předcházející části, v rámci služby **RCLIS** (Research in Computing, Library and Information Science, <http://rclis.org/>). Služba DoIS je řízena, spravována a udržována mezinárodní skupinou dobrovolníků z různých částí světa. Jde o univerzitní knihovníky, počítačové odborníky a učitele univerzit (také fakult Informační vědy) ze Španělska, Itálie, USA, Indie, Bosny a Hercegoviny aj. Elektronická knihovna je celosvětově volně dostupná a hostuje na serveru

Univerzity v Manchesteru (Velká Británie, <http://dois.mimas.ac.uk/>). Zrcadlové sídlo DoIS je lokalizováno v Itálii (<http://www.dois.it/>).

Hlavní úsilí mezinárodního týmu je v tuto chvíli vkládáno do tvorby **kvalitních metadat**, především metadat bibliografických jednotek, ukládaných dle stanovené metodiky nejdříve v lokálních archivech registrovaných systémem DoIS. Systém je tedy postaven na základě distribuované architektury (viz obr. č. 10). Metadata, která vytvářejí a ukládají výše zmiňovaní dobrovolníci, jsou z lokálních archivů sklízena pomocí harvestru a na bázi specifického Guildfordova protokolu (Guildford protocol) jsou ukládána v centrální databázi systému. Protokol připravil již v roce 1997 hlavní protagonista systému DoIS Thomas Krichel [KRICHEL, 1997a], který je jinak znám jako jeden z hlavních tvůrců úspěšného systému pro literaturu z ekonomického výzkumu RePEC. Bylo to v době, kdy nebyl ještě k dispozici protokol OAI-PMH. V současné době jsou však v rámci systému DoIS konány přípravy i na příjem metadat ze systémů archivů elektronických tisků, které jsou schopny předávat metadata na bázi protokolu OAI-PMH. Na tuto komunikaci je právě připravován spřízněný archiv E-LIS (viz obr. č. 10). A nic nebrání tomu, aby metadata nebyla sklízena i z oborových archivů dalších (například z archivu @rchiveSIC, jak je znázorněno na obrázku č. 10, nebo archivů dalších). Vzhledem k tomu, že samotný systém DoIS na protokol OAI-PMH není připraven, předpokládá se prozatím konverze metadat z formátu XML/Dublinské jádro do formátu systému DoIS, jímž je formát ReDIF (viz dále v textu).



Obr. č. 10: schéma současné (vlevo) a nově plánované (vpravo) architektury systému DoIS [upraveno, zdroj: BARRUECO, 2003b]

K 15. lednu 2004 obsahovala databáze systému DoIS 11 387 záznamů článků z elektronických časopisů a 3302 záznamů příspěvků z odborných akcí (konferencí, zasedání aj.). Zhruba polovina uložených záznamů (7653) je doplněna hyperodkazem na úplný text dokumentu. Ty jsou uloženy distribuovaně na různých serverech v rámci sítě Internet.

V rámci systému se intenzivně provádí i retrospektivní zpracování dokumentů. Pokud jde o konferenční sborníky, plánuje se další rozšíření. Zpracovávány by měly být nové tituly sborníků z významných akcí řady mezinárodních organizací.

Tým dobrovolníků systému DoIS doposud zjišťoval a zpracovával dva základní **typy dokumentů** z oboru informační věda: 1. časopisy a 2. sborníky z odborných akcí. V současné chvíli jsou již také připraveny specifikace pro popis knižních publikací a softwaru. **Popisnou jednotkou** je v převážné míře analytikum, tj. článek z časopisu a příspěvek ze sborníku z akce. U plánovaných knižních titulů se budou popisovat i jejich kapitoly. Záměrem správců systému je celosvětové pokrytí, nicméně již z letného pohledu na seznam pramenů je patrné, že čteně jsou zastoupeny prameny ze Španělska a Itálie (zejména ty, které nabízejí volný přístup k úplným textům), což je dáno také složením týmu dobrovolníků. K jádrovým excerpovaným pramenům patří známé, hojně využívané a volně dostupné časopisy, například časopis „Bulletin des Bibliothèques de France“ (cca 450 záznamů), „Ariadne“ (cca 370 záznamů), „D-Lib Magazine“ (cca 300 záznamů) aj. Nicméně, náhled do rejstříku excerpovaných časopisů prozrazuje, že velká množství záznamů reprezentují i články z komerčně dostupných časopisů. Absolutně největší počet záznamů pokrývá produkci prestižního časopisu JASIST, který vydává americká společnost ASIST. Počet excerpovaných záznamů z časopisu JASIST dosahuje počtu téměř 1000 jednotek.

**Template-Type:** ReDIF-Article 1.0

**Author-Name:** Schwarz, Josef

**Author-Email:** schwarzjv@seznam.cz

**Author-Homepage:** <http://schwarz.webpark.cz/>

**Title:** Vybrané teoretické a metodologické problémy terminografie : poznatky z tvorby České terminologické databáze knihovnictví a informační vědy = Selected theoretical and methodological problems of terminography: knowledge acquired in building the Czech terminological database of library and information science

**Abstract:** An article is based on project called "Czech terminological database of library and information science (TDLIS)" that was finished last year (2002). The results of the project are described from theoretical, methodological and historical perspectives. Basic terminographical concepts and theories are shortly outlined, methods and techniques of terminological database compilation are discussed in more details. Key phases are: terminographical team formation, definition of basic project starting points, excerption and identification of concepts and terms and terminological entries compilation. Each step is introduced by description of general methods and concrete examples of TDLIS project solutions are added. History of the Czech and Slovak terminography after 1945 is mentioned and basic information about TDLIS are presented along with overview of TDLIS future developments.

**Keywords:** Library and information science; Terminology management; Terminology; Terminography; Terminological database; Theory; Methodology; History; Czech Republic; Slovak Republic

**Journal:** Národní knihovna : knihovnická revue

**Pages:** 21-41

**Volume:** 13

**Issue:** 1

**Year:** 2003

**File-URL:** <http://full.nkp.cz/nkkz/Nkkz0204/0204268.html>

**File-Format:** text/html

**File-URL:** <http://full.nkp.cz/nkkz/pdf/0301/0301021.pdf>

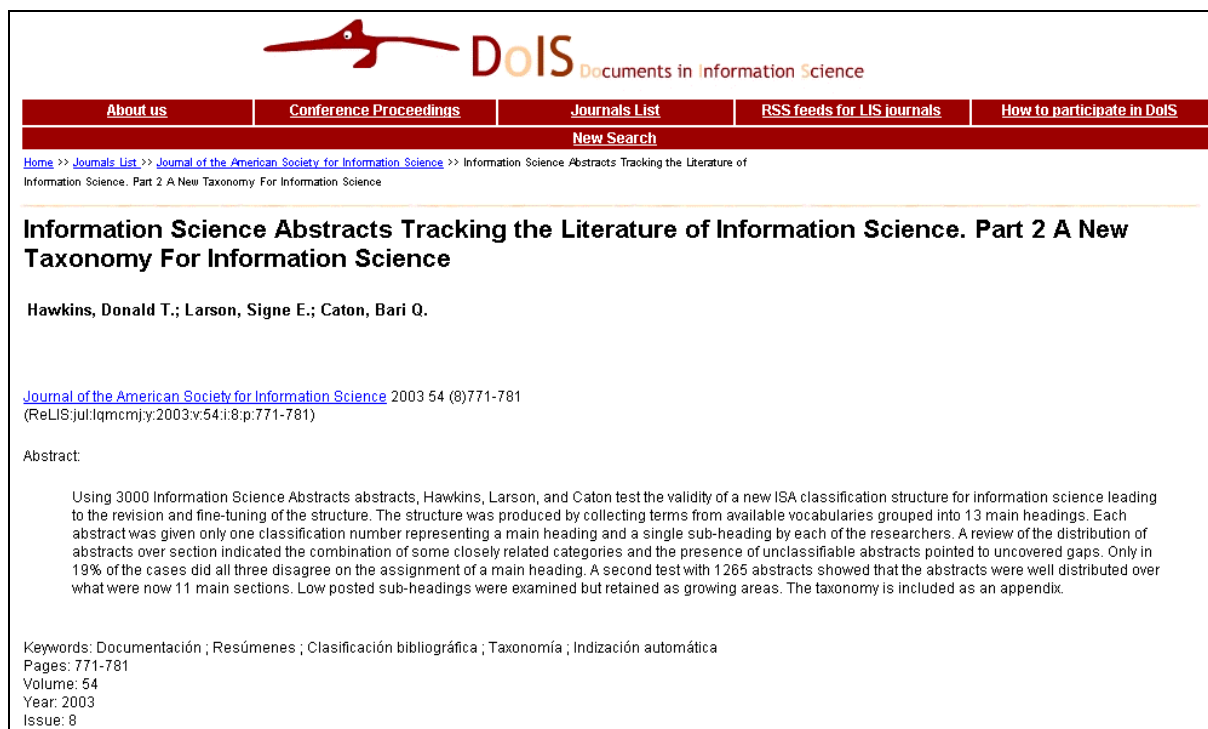
**File-Format:** application/pdf

**Handle:** ReLIS:xxx:nkkcz:v:13:y:2003:i:1:p:21-41

*Obr. č. 11: příkladový záznam článku z českého časopisu ve formátu ReDIF (záznam není součástí báze DoIS, jeho identifikátor Handle není reálný) [údaje pocházejí z časopisu Národní knihovna]*

Záznamy dokumentů jsou připravovány dle specifikace již výše zmíněného formátu metadat **ReDIF** (Research Documentation Information Format), která vychází z definic známých formulářů IAFA (IAFA templates). Specifikace metadat ReDIF byla původně připravena v rámci výše zmíněného systému pro výzkumné práce a literaturu z oblasti ekonomie RePEc (Research Papers in Economics, <http://repec.org/>), který je jedním z významných účastníků iniciativy propojování elektronických archivů OAI [KRICHEL, 1997b]. V záznamech metadat je, vedle jiných údajů (velmi důležitý a cenný je abstrakt v anglickém jazyce), podstatný identifikátor (viz obr. č. 11 a 12), který je modelován v souladu s požadavky

systému DOI (<http://www.doi.org>). Záznam může také obsahovat odkaz vedoucí k získání plných textů, pokud ovšem nejde o případ komerčně dostupného zdroje (pak adresa na plný text chybí). Na obr. č. 11 je pro představu uveden záznam článku z českého časopisu Národní knihovna ve formátu ReDIF, tedy ve struktuře potřebné v tento okamžik k uložení v bázi DoIS (nejde tedy o záznam v bázi fakticky uložený). Bylo by vítáno, kdyby se záznamy vybraných článků z tohoto časopisu mohly v blízké budoucnosti automaticky dostávat do báze DoIS.



**DoIS** Documents in Information Science

[About us](#) | [Conference Proceedings](#) | [Journals List](#) | [RSS feeds for LIS journals](#) | [How to participate in DoIS](#)

[New Search](#)

[Home](#) >> [Journals List](#) >> [Journal of the American Society for Information Science](#) >> Information Science Abstracts Tracking the Literature of Information Science. Part 2 A New Taxonomy For Information Science

**Information Science Abstracts Tracking the Literature of Information Science. Part 2 A New Taxonomy For Information Science**

Hawkins, Donald T.; Larson, Signe E.; Caton, Bari Q.

[Journal of the American Society for Information Science](#) 2003 54 (8)771-781  
(ReLIS:jul:qmcmy:2003:v:54:i:8:p:771-781)

Abstract:

Using 3000 Information Science Abstracts abstracts, Hawkins, Larson, and Caton test the validity of a new ISA classification structure for information science leading to the revision and fine-tuning of the structure. The structure was produced by collecting terms from available vocabularies grouped into 13 main headings. Each abstract was given only one classification number representing a main heading and a single sub-heading by each of the researchers. A review of the distribution of abstracts over section indicated the combination of some closely related categories and the presence of unclassifiable abstracts pointed to uncovered gaps. Only in 19% of the cases did all three disagree on the assignment of a main heading. A second test with 1265 abstracts showed that the abstracts were well distributed over what were now 11 main sections. Low posted sub-headings were examined but retained as growing areas. The taxonomy is included as an appendix.

Keywords: Documentación ; Resúmenes ; Clasificación bibliográfica ; Taxonomía ; Indización automática

Pages: 771-781  
Volume: 54  
Year: 2003  
Issue: 8

Obr. č. 12: výstupní metadatový záznam článku z komerčního časopisu JASIS uložený v bázi systému DoIS [převzato 2004-01-15]

Reálný záznam metadat článku publikovaného v časopise JASIST je uveden na obr. č. 12. Byl připraven v jednom z lokálních archivů na serveru Univerzity v Salamance (Španělsko), jehož identifikátor je „RCLIS:jul“. Záznamy pro archiv připravuje jeden z dobrovolníků systému DoIS Julio Alonso Arévalo z Univerzity v Salamance [BARRUECO CRUZ, 2003b, část 2.1]. Po bližším prozkoumání údajů záznamů a jejich srovnání s údaji uvedenými na stránkách časopisu JASIST (<http://www.asis.org/Publications/JASIS/index.html>) je zřejmé, že je zcela využito údajů z amerického serveru, klíčová slova ve španělštině jsou doplněna ze strany zpracovatele. Zpracovatel pravděpodobně využívá nějakého vlastního programového nástroje na převod základních metadat, nicméně jisté ruční zásahy jsou asi nutné, a proto je třeba tuto práci zvláště ocenit. Pokud by se v blízké budoucnosti začaly budovat na úrovni sídel významných komerčních časopisů volně dostupné databáze metadat popisujících jejich články (nejlépe v souladu s již zmiňovaným protokolem OAI-PMH), mohla by být práce systému DoIS, popřípadě jakéhokoliv dalšího systému, bezproblémová. A bylo by to ku prospěchu jak koncových uživatelů, tak komerčního uživatele samotného. Lze hovořit i o tom, že začíná nová éra mezinárodní oborové bibliografie. Služba DoIS k takovému produktu směřuje, je dostupná zdarma a lze doufat, že i nadále zůstane. Je ale nutné hledat způsoby, jak do tohoto systému postupně zapojit více aktivních účastníků. Stávající tým proto ve své dokumentaci k systému vyzývá k celosvětové spolupráci. Jednou z možností je i zapojení se do archivu elektronických tisků E-LIS, který byl založen v rámci služby DoIS a bude prvním archivem,

jehož metadata budou sklížena do databáze DoIS na základě protokolu OAI-PMH (proces se právě připravuje). Porovnáním záznamu článku na obr. 12 se záznamem uvedeným v sídle oficiálního nakladatele časopisu JASIST (Wiley) v digitální knihovně InterScience lze zjistit, že se liší v podstatě jenom na úrovni abstraktu (abstrakt v komerční digitální knihovně je obsáhlejší).

**Vyhledávání informací** z databáze metadat systému DoIS má prozatím jednoduchou povahu. Prvním způsobem vyhledávání je standardní jednoduché **prohlížení rejstříků/seznamů**: 1. zdrojových (excerpovaných) elektronických časopisů a 2. zdrojových (excerpovaných) konferenčních sborníků. V seznamech je v tomto případě je dobře vidět, kde jsou a kde nejsou dostupné plné texty.

Druhým způsobem je **přímé formulování rešeršních dotazů** pomocí WWW formuláře rozhraní DoIS. Dotazy je možné formulovat pomocí klíčových slov, která budou prohledávána v celém záznamu metadat. Systém nabízí i hledání v polích Autor (Author) a Název (Title) a Popis (Description). Dotazovací jazyk umožňuje využívání základních booleovských operátorů a možnost rozšíření selekčních údajů pomocí znaku „\*“.

Specifickou novou službou systému DoIS je signální služba **RSS** (Rich Site Summary), na základě které jsou uživatelé zpravováni automaticky o obsazích nových čísel volně dostupných elektronických časopisů v momentě, kdy se objeví na WWW. Prozatím jsou nabízeny články ze čtyř známých časopisů z oboru informační věda: amerického časopisu „D-Lib Magazine“ a tří britských časopisů „Ariadne“, „Information Research“ a „Journal of Digital Information“. Záznamy jsou komunikovány na bázi formátu XML.

## 5 Závěry

Do světa informačních zdrojů v oboru informační věda vstoupily s intenzivním rozvojem Internetu a informačních technologií dva jejich nové typy: archivy elektronických tisků, které se pokoušejí na jednom místě shromažďovat digitální dokumenty vědeckovýzkumného charakteru, a informační služby hybridního charakteru, jejichž prvořadým úkolem je globální bibliografické mapování oboru informační věda, tedy budování novodobé oborové bibliografie obohacené o odkazy na dostupné úplné texty dokumentů. Oba nové typy silně zasahují a dá se říci i postupně proměňují dosavadní tradiční strukturu informačních zdrojů. Archivy elektronických tisků lze pokládat za novodobé seriály, resp. integrující zdroje, jak jsou dle nejnovější terminologie označovány. V tuto chvíli existují v oboru informační věda již tři archivy a zatím si lze jenom v obrysech představit, kolik jich může přibýt během dalších let, jak budou konkrétně profilovány, jaká bude jejich celková skladba a jak ovlivní komunikaci v oboru.

Analýza čtyř zajímavých systémů ukázala, že v souvislosti s tím, jak se rozvíjely informační technologie, postupně vznikaly jejich jednotlivé typy. Historicky prvním systémem byla služba DoIS, jejíž představitelé se na bázi zcela dobrovolné pokoušejí vytvořit konkurenci tradičním komerčním bibliografickým a jiným službám z oboru informační věda. Hlavním motivem však nebyla ani tak touha konkurovat, jako využít té situace a skutečnosti, že se v prostředí WWW během desetiletí nakupilo již značné množství literatury, jejíž nezanedbatelná část je volně dostupná. Konečně, komerční bibliografické služby již také delší dobu zpracovávají i volně dostupné dokumenty z Internetu, avšak zpřístupňují je za úplaty, což z jistého pohledu vypadá jako paradox. Bylo by nesmírně užitečné, kdyby i v oboru informační věda vznikla volně dostupná bibliografická služba, která by globálně pokrývala alespoň všechny podstatné volně dostupné i komerční časopisy a periodické sborníky z oboru. Stačilo by k tomu teoreticky málo: opatřit databáze, na základě nichž i komerční vydavatelé zdarma (!) prezentují prostřednictvím záznamů obsahy svých časopisů, nezbytným minimem

potřebných technologií a protokolů (OAI-PMH aj.) tak, aby všechny libovolné služby mohly automaticky sklízet metadata a budovat z nich potřebné bibliografické výstupy. Dobrovolníkům ze systému DoIS náleží v tomto prvenství a bylo by velice vhodné tuto službu do budoucna nejen propagovat, ale i napomáhat jejímu zdárnému rozvoji ve prospěch všech uživatelů.

Rok 2000 byl přelomovým rokem v rozvoji tzv. archivů elektronických tisků. Dočkal se i obor informační věda, a tak v prostředí k tomu nejvhodnějším, tedy akademicko-univerzitním, vznikla trojice těchto archivů. Značný vliv na jejich rozvoj měl jednak vznik iniciativy OAI a jejího hlavního produktu - protokolu pro automatický sběr metadat OAI-PMH a budování návazných služeb, jednak intenzivní rozvoj návazných technologií a volně dostupných softwarových produktů pro budování těchto archivů (Eprints.org, DSpace aj.). Dispozice programů a technologií je ovšem jenom předpoklad a počátek budování archivů volně dostupných digitálních dokumentů z výzkumů z oboru informační věda. Nezbytné je archiv jistým způsobem řídit a dále udržovat. Tři analyzované archivy dokladují, že schopný a erudovaný personál, který by se o archivy staral, existuje (jistá dávka nadšení nikde nechyběla). Dá se říci, že provozy těchto archivů jsou novým prvkem celkové informační činnosti univerzit, které by tudíž měly najít minimální prostředky na jejich provoz. Ukazuje se, že se dobře a úspěšně dají takové archivy budovat i v rámci mezinárodní spolupráce (archiv E-LIS).

Ze tří analyzovaných archivů se úspěšně rozvíjejí především evropské projekty. Zajímavý a svěbytný francouzský výzkum věd o informacích je pozoruhodně prezentován v archivu @rchiveSIC. Jeho obsah je vysoce kvalitní, bohužel ne všechny příspěvky však bude moci využívat větší počet uživatelů (jazykové bariéry). Nicméně, v archivu se nalézají i dokumenty v anglickém jazyce a podstatné příspěvky francouzských odborníků zaznívají také v angličtině na různých fórech. A tak se pravděpodobně budou dostávat i do druhého evropského archivu E-LIS, který v poslední době značně rozšiřuje objem svého repozitáře i díky iniciativní excerpci příspěvků z řady mezinárodních konferencí konaných na různých místech Evropy i jinde. Pozadu zatím zůstává archiv americký. Pokud se mu však podaří nově koncipovaný záměr rozšíření stávajícího systému DLIST dovést do úspěšného konce, můžeme se dočkat zajímavého digitálního fondu, který bude sloužit v oblasti informační a technologické gramotnosti.

Všechny analyzované archivy úspěšně rozvíjejí dílčí metodické otázky, jako jsou metadatové popisy dokumentů nebo nabízené služby. Daří se rozvíjet další nadstavbové služby, jako je například propojování citovaných a zdrojových dokumentů a budoucí budování citačních rejstříků vystavených nad archivy. Zajímavým problémem archivů z oboru informační věda je věcné pořádání informací. Všechny analyzované archivy přistoupily minimálně k jednoduchému schématu předmětových kategorií, ve kterých se promítá základní problematika i koncepce oboru informační věda ve vybraných zemích či regionech. Ukazuje se, že předmětové třídění je v archivech nezbytné jak pro generování předmětového rejstříku, který uživatelé rádi navštěvují, tak například pro realizaci signálních služeb (typu Alert). Na příkladu služby DoIS, která nemá zavedeno podobné třídění, je vidět značné ochuzení uživatelů v procesech vyhledávání informací.

Specifickou otázkou, která se debatuje, a to nejen v oblasti archivů elektronických tisků, je otázka autorskoprávního charakteru. Zdá se, že se jí daří překonávat a že komerční vydavatelsko-distributorská sféra se snaží s novou situací vyrovnat. Hledá nové směry a způsoby své existence. Díky novým technologiím se ale v univerzitním prostředí daří konečně budovat novodobé digitální fondy jejich vlastní produkce, jež jsou celosvětově dostupné všem zájemcům, kteří tak příslušné obory mohou rychleji a efektivněji posunovat dále.



## Použité informační zdroje

- ❑ BARRUECO CRUZ, J.M.; KLINK, M.J.R.; KRICHEL, T. 2003a. Personal Data in a large digital library. In *Proceedings Fourth European Conference on Research and Advanced Technology for Digital Libraries, Lisbon (Portugal)*. 2003. Preprint dostupný také z E-LIS: <<http://eprints.rclis.org/archive/00000435/01/phoenix.a4.pdf>>
- ❑ BARRUECO CRUZ, J.M.; SUBIRATS COLL, I. 2003b. CRLIS : towards a digital library for Information Science. In *Proceedings Libraries in Digital Age (LIDA), Dubrovnik and Mljet (Croatia)*. 2003-11-27. Preprint dostupný také z E-LIS: <<http://eprints.rclis.org/archive/00000352/01/LIDA.pdf>>. PPT prezentace dostupná také z: <<http://eprints.rclis.org/archive/00000352/02/subirats.ppt>>
- ❑ BARRUECO CRUZ, J.M.; SUBIRATS COLL, I. 2003c. Documents in Information Science : A free space for our profession on the Internet. *Serials : The journal for the serials community*. 2003, vol. 16, issue 2, s. 211-214. ISSN 0953-0460. Dostupný také z WWW: <<http://www.sparcinfo.blogspot.com/cruz.doc>>
- ❑ BRATKOVÁ, E. 2000. Elektronické archívy vědecké literatury a jejich integrace. In *INFOS 2000 : zborník z 30. medzinárodného informatického sympózia , ktoré sa konalo v dňoch 3.-6. apríla 2000 v Starej Lesnej*. Bratislava : Spolok slovenských knihovníkov, 2000, s. 187-200. Dostupný také z WWW: <<http://www.aib.sk/infos/infos2000/19.htm>>. ISBN 80-81565-80-5.
- ❑ BRATKOVÁ, E. 2002. Elektronické disertace a projekty a systémy jejich zpřístupňování. *Národní knihovna* [online]. 2002, roč. 13, č. 4, s. 268-287 [2004-01-15]. Dostupný z WWW: <<http://full.nkp.cz/nkkr/Nkkr0204/0204268.html>>. ISSN 1214-0678.
- ❑ CARPENTER, L. 2001. *OAI for Beginners - The Open Archives Forum online tutorial* [online]. [S.l.] : Open Archives Forum, c2001- , last modif. 2003-10-14 [cit. 2004-01-15]. Dostupný z WWW: <<http://www.oaforum.org/tutorial/>>
- ❑ COLEMAN, A.; BRACKE, P. 2003a. *DLIST : Building An International Scholarly Communication Consortium for Library and Information Science* [online]. 2003 [cit. 2004-01-15]. Dostupný z WWW: <<http://dlist.sir.arizona.edu/pubs/6.htm>>
- ❑ COLEMAN, A.; MALONE, Ch.; BRACKE, P. 2003b. *DLIST : Digital Library of Information Science & Information Science* [online prezentace]. 2003-03-18 [cit. 2004-01-15]. 19 snímků. Dostupný z WWW: <<http://www.sir.arizona.edu/faculty/coleman/ualibrary.ppt>>
- ❑ COLEMAN, A.; MALONE, Ch.; BRACKE, P. 2003c. *DLISTConnection : Information and Technology Literacy Service for NSDL* [online]. 21 October 2003 [cit. 2004-01-15]. Dostupný z DLIST: <<http://dlist.sir.arizona.edu/archive/00000271/01/dlistconnection.pdf>>
- ❑ De ROBBIO, A. 2003. E-LIS: un Open Archive in Library and Information Science. *Bibliotime : Rivista elettronica per le biblioteche* [online]. Marzo 2003, anno 6, no. 1 [cit. 2004-01-15]. Dostupný z WWW: <<http://www.spbo.unibo.it/bibliotime/num-vi-1/derobbio.htm>>. ISSN 1128-3564. Dostupný také z E-LIS: <<http://eprints.rclis.org/archive/00000201/01/E-LIS-bibtime.pdf>>
- ❑ Eprints.org. 2002. *Eprints.org. Self-Archivig FAQ* [online]. Southampton : University of Southampton, c2002 [cit. 2004-01-15]. Dostupný z WWW. <<http://www.eprints.org/self-faq/>>
- ❑ GALLEZOT, G. 2003. *ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la Communication : rôle, fonctionnement et usage* [online]. 18 September 2003 [cit. 2004-01-15]. Publikováno v příručce Guide pratique : Les publications électroniques (Archimag, 2003). Dostupný z @rchive SIC: <[http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/06/00/sic\\_00000600\\_01/sic\\_00000600.pdf](http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/06/00/sic_00000600_01/sic_00000600.pdf)>

- ❑ GALLEZOT, G.; CHARTRON, G.; NOYER, J.-M. 2002. *Une archive ouverte des publications en InfoCom* [online]. 17 May 2002 [cit. 2004-01-15]. Preprint z kolokvia „Place et enjeux des revues pour la recherche en infoCom (SFSIC)“, 25 March 2002. Dostupný z @rchive SIC: <[http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/00/29/sic\\_00000029\\_00/sic\\_00000025.pdf](http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/00/29/sic_00000029_00/sic_00000025.pdf)>
- ❑ HARNAD, S. 2001. The self-archiving initiative : Freeing the refereed research literature online. *Nature : Web Debates* [online]. 26 April 2001, no. 410, s. 1024-25 [cit. 2004-01-15]. Dostupný z WWW: <<http://www.nature.com/nature/debates/e-access/Articles/harnad.html>>. Dostupný také z CogPrints: <<http://cogprints.ecs.soton.ac.uk/archive/00001642/01/nature4.htm>>
- ❑ HAWKINS, D.T.; LARSON, S.E.; CATON, B.Q. 2003. Information Science Abstracts : Tracking the Literature of Information Science. Part 2, A New Taxonomy for Information Science. *Journal of the American Society for Information Science and Technology*. 2003, vol. 54, no. 8, s. 771-781. Dostupný z také Wiley InterScience: <<http://www3.interscience.wiley.com/cgi-bin/home>>
- ❑ KRICHEL, T. 1997a. *Guildford Protocol* [online]. [1997] [cit. 2004-01-15]. Dostupný z WWW: <<ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/guilp.html>>
- ❑ KRICHEL, T. aj. 1997b. *ReDIF* [online]. Version 1. [1997] [cit. 2004-01-15]. Dostupný z WWW: <[ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/redif\\_1.html](ftp://netec.mcc.ac.uk/pub/NetEc/RePEc/all/root/docu/redif_1.html)>
- ❑ LAGOZE, C. aj. 2003. *The Open Archives Initiative Protocol for Metadata Harvesting* [online]. Protocol Version 2.0 of 2002-06-14, Document version 2003/02/21T00:00:00Z. 2003 [cit. 2004-01-15]. Dostupný z WWW: <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>